



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2020

---

## **Terrifying film music mimics alarming acoustic feature of human screams**

Trevor, Caitlyn ; Arnal, Luc H ; Frühholz, Sascha

**Abstract:** One way music is thought to convey emotion is by mimicking acoustic features of affective human vocalizations [Juslin and Laukka (2003). *Psychol. Bull.* 129(5), 770-814]. Regarding fear, it has been informally noted that music for scary scenes in films frequently exhibits a "scream-like" character. Here, this proposition is formally tested. This paper reports acoustic analyses for four categories of audio stimuli: screams, non-screaming vocalizations, scream-like music, and non-scream-like music. Valence and arousal ratings were also collected. Results support the hypothesis that a key feature of human screams (roughness) is imitated by scream-like music and could potentially signal danger through both music and the voice.

DOI: <https://doi.org/10.1121/10.0001459>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-191635>

Journal Article

Published Version

Originally published at:

Trevor, Caitlyn; Arnal, Luc H; Frühholz, Sascha (2020). Terrifying film music mimics alarming acoustic feature of human screams. *Journal of the Acoustical Society of America*, 147(6):EL540-EL545.

DOI: <https://doi.org/10.1121/10.0001459>

## Terrifying film music mimics alarming acoustic feature of human screams

Caitlyn Trevor, Luc H. Arnal, and Sascha Frühholz

Citation: [The Journal of the Acoustical Society of America](#) **147**, EL540 (2020); doi: 10.1121/10.0001459

View online: <https://doi.org/10.1121/10.0001459>

View Table of Contents: <https://asa.scitation.org/toc/jas/147/6>

Published by the [Acoustical Society of America](#)

---

### ARTICLES YOU MAY BE INTERESTED IN

[Change in acoustic impulse response of a room due to a fire](#)

[The Journal of the Acoustical Society of America](#) **147**, EL546 (2020); <https://doi.org/10.1121/10.0001415>

[It is too loud!](#)

[The Journal of the Acoustical Society of America](#) **148**, R3 (2020); <https://doi.org/10.1121/10.0001666>

[Intelligibility of chimeric locally time-reversed speech](#)

[The Journal of the Acoustical Society of America](#) **147**, EL523 (2020); <https://doi.org/10.1121/10.0001414>

[Effect of instrument timbre on musical emotion recognition in normal-hearing listeners and cochlear implant users](#)

[The Journal of the Acoustical Society of America](#) **147**, EL535 (2020); <https://doi.org/10.1121/10.0001475>

[Machine learning in acoustics: Theory and applications](#)

[The Journal of the Acoustical Society of America](#) **146**, 3590 (2019); <https://doi.org/10.1121/1.5133944>

[Perception of musical melody and rhythm as influenced by native language experience](#)

[The Journal of the Acoustical Society of America](#) **147**, EL385 (2020); <https://doi.org/10.1121/10.0001179>

---

Read Now

**JASA**  
THE JOURNAL OF THE  
ACOUSTICAL SOCIETY OF AMERICA

**Special Issue:** Phonetics of  
Under-Documented Languages

# Terrifying film music mimics alarming acoustic feature of human screams

.....  
Caitlyn Trevor,<sup>1</sup> Luc H. Arnal,<sup>2</sup> and Sascha Frühholz<sup>1,a)</sup>

<sup>1</sup>Department of Psychology, University of Zurich, Binzmuehlestrasse 14, 8050 Zurich, Switzerland  
<sup>2</sup>Department of Fundamental Neuroscience, University of Geneva, Biotech Campus, Geneva 7, CH-1202, Switzerland  
[caitlyn.trevor@psychologie.uzh.ch](mailto:caitlyn.trevor@psychologie.uzh.ch), [luc.arnal@unige.ch](mailto:luc.arnal@unige.ch), [sascha.fruhholz@psychologie.uzh.ch](mailto:sascha.fruhholz@psychologie.uzh.ch)

**Abstract:** One way music is thought to convey emotion is by mimicking acoustic features of affective human vocalizations [Juslin and Laukka (2003). *Psychol. Bull.* **129**(5), 770–814]. Regarding fear, it has been informally noted that music for scary scenes in films frequently exhibits a “scream-like” character. Here, this proposition is formally tested. This paper reports acoustic analyses for four categories of audio stimuli: screams, non-screaming vocalizations, scream-like music, and non-scream-like music. Valence and arousal ratings were also collected. Results support the hypothesis that a key feature of human screams (roughness) is imitated by scream-like music and could potentially signal danger through both music and the voice. © 2020 Acoustical Society of America

[Editor: Brenda L. Lonsbury-Martin]

Pages: EL540–EL545

Received: 25 March 2020 Accepted: 2 June 2020 Published Online: 24 June 2020

## 1. Introduction

Music used to underscore frightening scenes in movies is often described as sounding “scream-like.” A well-known example is the music accompanying the infamous shower murder scene in Alfred Hitchcock’s film *Psycho* (1960) with “screeching, upward glissandi” from the violins [Brown (1982), p. 46]. Although “scream-like” is a common descriptor, the question remains: do these scary film soundtrack excerpts *actually* sound like and are perceived similarly to human screams?

Music has a long history of relying on vocal behaviors to describe musical ones. For example, a branch of music theory (topic theory) has a categorical label (*pianto*) to describe music that mimics the sound of human weeping or sighs (Mirka, 2014). Recently, some music cognition researchers have begun to empirically investigate such instances of mimicry (Huron and Trevor, 2017; Trevor and Huron, 2019). These investigations are part of a branch of music and emotion research that theorizes that music might sometimes communicate emotion by mimicking human ethological vocal signals (Juslin and Laukka, 2003; Blumstein *et al.*, 2012; Bryant, 2013; Huron, 2015; Warrenburg, 2019). Ethological signals are behaviors intended to communicate with a fellow member of one’s species and cause them to react in a desired manner (Ehret, 2006; Lorenz, 1939). In humans, ethological signals can be smiling, crying, screaming, etc. (Huron, 2015; Ohala, 1996). Inspired by this branch of research and by the frequent comparison of scary music to human screams, the motivating question for the current study is: Do scream-like musical passages in scary film music actually mimic the sound of human screams to scare viewers?

What acoustic features characterize the sound of a human scream? Typically, human screams are loud, utilize a wide range of frequencies, are higher in pitch than one’s average vocal range, and have a high amount of roughness (Arnal *et al.*, 2015; Schwartz *et al.*, 2019). Roughness is a basic auditory phenomenon that is characterized by a coarse, grating, or harsh subjective experience (Terhardt, 1974). Roughness has been defined in various ways, and a number of models and operationalizations have been formulated (Vassilakis and Kendall, 2010). Most models focus on amplitude modulation, especially in the broad region between 15 and 200 Hz. For the purposes of this study, we employ the modulation power spectrum (MPS) method and parameters used by Arnal *et al.* (2015). The MPS is a two-dimensional Fourier transformation of a soundwave that quantifies both temporal and spectral power modulations (Elliott and Theunissen, 2009). Previous research indicates that human screams feature higher MPS values than non-alarming vocalizations in the 30 to 150 Hz range of the temporal modulation rate dimension of the MPS (Arnal *et al.*, 2015) [see Fig. 1(A)]. Vocal and artificial sounds exhibiting high temporal modulations in this range are perceived as particularly aversive (Li

<sup>a)</sup>Also at: Department of Psychology, University of Oslo, Forskningsveien 3A, 0373 Oslo, Norway.

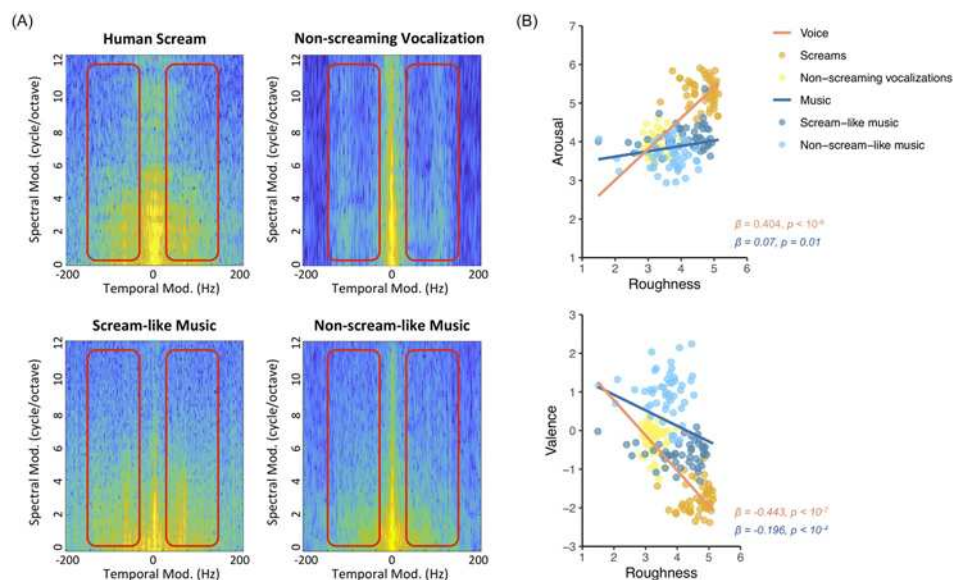


Fig. 1. (Color online) (A) Presented here are four MPS analyses of four audio recordings (from top left moving clockwise): a human scream, a non-screaming human vocalization (a held “ah” sound), a non-scream-like film music excerpt (T41), and a scream-like film music excerpt (T04) [music excerpts available on OSF, [Trevor et al. \(2020\)](#)]. The MPS is a two-dimensional Fourier transformation of a soundwave that quantifies both temporal and spectral power modulations ([Elliott and Theunissen, 2009](#)). The circles outline the temporal modulation range of 30 to 150 Hz. Sounds that exhibit roughness (coarse, grating sounding) ([Terhardt, 1974](#)), such as alarming sounds like fire alarms and human screams, have been shown to activate that range of the MPS ([Arnal et al., 2015](#)). In the figure, it is evident that the human scream and the scream-like music exhibit strong activation patterns in the roughness region of the MPS while the non-screaming vocalization and the non-scream-like musical excerpt exhibit much weaker ones. (B) The upper plot shows the relationship between the average arousal ratings and average roughness of the four stimulus categories. There is a significant positive correlation between arousal and roughness for both vocal ( $\beta = 0.404, p < 10^{-6}$ ) and musical ( $\beta = 0.07, p = 0.01$ ) stimuli. However, the correlation is significantly more pronounced for vocal stimuli as opposed to musical stimuli ( $p < 10^{-5}$ ). The lower plot shows the relationship between the average valence ratings and average roughness of the four stimulus categories. There is a significant negative correlation between valence and roughness for both vocal ( $\beta = -0.443, p < 10^{-7}$ ) and musical ( $\beta = -0.196, p < 10^{-4}$ ) stimuli. However, the correlation is significantly more pronounced for vocal stimuli as opposed to musical stimuli ( $p < 10^{-5}$ ).

[et al., 2018](#)), cause faster behavioral reactions ([Arnal et al., 2015](#); [Ollivier et al., 2019](#)), and increase neural responses in subcortical brain regions associated with aversive processing ([Arnal et al., 2019](#)).

The other features of screams (high intensity, broad spectrum, and high pitch) raise thorny measurement issues. Sound pressure levels cannot be measured directly from recorded audio files. High pitch can be gauged only with respect to a speaker’s normative vocal register. Broad spectrum is evident in a wide range of vocalizations. Laughter, for example, can also feature high loudness, wide spectral range, and high relative pitch. It is the relative uniqueness and ease of measurement that makes roughness a useful operational measure, and coincidentally, a prime candidate for a universal cue signaling danger ([Arnal et al., 2015](#)).

To investigate whether scream-like music has the same roughness feature as, and is perceived similarly to, human screams, we conducted two studies. In the first study, we ran an acoustic analysis to test whether recorded screams and scream-like music exhibit enhanced roughness compared with control recordings. In the second study, we collected valence and arousal ratings for the audio files in order to test whether screams and scream-like music are perceived as sharing similar emotional qualities. We made the following hypotheses. First, we hypothesized that the mean power of the MPS within the roughness region (henceforth “roughness”) would be similar for screams and scream-like music, and would be significantly greater for screams compared to non-screaming vocalizations and for scream-like music compared to non-scream-like music. Second, given that roughness may be a universal cue for danger ([Arnal et al., 2015](#)), we hypothesized that roughness would correlate negatively with valence ratings and positively with arousal ratings for both music and vocal stimuli. Taken together, these results would demonstrate that scream-like music both sounds like and is perceived similarly to actual human screams.

## 2. Methods

### 2.1 Stimuli and collection of audio recordings

The audio recordings used in the studies were deployed in a  $2 \times 2$  factorial design, with one factor corresponding to the sound *source* (music, voice) and the other one to the *scream-likeness* of



the sounds (scream-like, non-scream-like). Specifically, the four collections included (a) fearful scream vocalizations, (b) scream-like film music excerpts, (c) non-fearful human vocalizations (sounding similar to a held “ah” sound), and (d) non-scream-like film music excerpts as controls. All audio recordings are 800 ms in duration, RMS normalized, sampled at 16 kHz, and are in wav-file format.

In assembling a database of scream-like and non-scream-like music, we chose to make use of excerpts from horror film soundtracks. We chose the horror genre because descriptions of scream-like music that we found were typically describing horror movie soundtracks [i.e., [Brown \(1982\)](#)] and because horror films soundtracks are written with the explicit aim of scaring viewers [as opposed to violent or aggressive music, like death metal, which has been found to induce a range of positive and negative emotions in listeners ([Thompson et al., 2019](#))]. Using an expertise-based approach, we curated excerpts from ten recently released films (2010 or later) that employed original composed soundtracks. Five scream-like music excerpts and five non-scream-like music excerpts were sampled from each of the ten horror movie soundtracks. In selecting potential scream-like passages, sampling focused on music written for especially terrifying scenes, such as moments of attack by a monster or ghost. On the other hand, non-scream-like music excerpts were pulled from scenes that were not especially terrifying and for which the music was deemed to be more emotionally neutral [see Table S1 for more information<sup>1</sup>; to download the excerpts see [Trevor et al. \(2020\)](#)]. The curation process resulted in a database of 50 scream-like excerpts and 50 non-scream-like excerpts (Table S2 identifies the full list of excerpts and films<sup>1</sup>).

Fearful screams and non-scream vocalizations were recorded from ten participants (five female, age:  $M = 27.20$ ,  $SD = 3.39$ ). Non-scream vocalizations consisted simply of a sustained “ah” vowel. Although the scream-like vocalizations were acted, past research has demonstrated that even acted screams have been judged to resemble real-life screams to a high degree ([Engelberg and Gouzoules, 2018](#)). Each participant produced ten audio recordings, resulting in a total of 50 fearful screams and 50 non-scream vocalizations.

## 2.2 The acoustic analyses

Using MATLAB, the MPS of each excerpt was measured using the same procedure and equations as used by Arnal and colleagues ([Arnal et al., 2015](#)). Specifically, the initial spectrograms were obtained using a filter-bank approach with 128 Gaussian windows whose outputs were Hilbert transformed and then log-transformed. Then the modulation power spectra were obtained by applying a two-dimensional Fourier transform to the spectrogram (24 channels/octave) and log-transforming the resulting spectral power density estimates ([Arnal et al., 2015](#)). From there, the mean amplitude in the roughness range of 30 to 150 Hz along the temporal modulation range was taken [see Fig. 1(A); see [Elliott and Theunissen \(2009\)](#) for more detailed information on the MPS]. Results and statistical tests are reported after describing the behavioral study.

## 2.3 Participants in ratings experiment

In the second study, 20 healthy participants (twelve female) reporting normal hearing and no psychiatric disorders participated in the rating experiment. Participants were recruited through the University of Zurich and received 15 CHF for their participation in the experiment. The participants were between the ages of 21 and 37 ( $M = 26.10$ ,  $SD = 4.02$ ). The procedure was approved by the Cantonal Ethics Commission of Zurich, Switzerland.

## 2.4 Experimental procedure of the rating experiment

The experimental task for the rating experiment consisted of listening to each of the 200 audio files and rating the valence (from  $-3$  to  $3$ , with “ $-3$ ” indicating the most negative valence and “ $3$ ” indicating the most positive valence) and arousal (from  $1$  to  $7$ , with “ $1$ ” indicating lowest arousal and “ $7$ ” indicating highest arousal) of the conveyed emotion using two analogical-categorical sliding scales. The experiment interface was created with MATLAB using Psychtoolbox. It took place in a quiet research room at the University of Zurich on a PC desktop computer using Sennheiser HD 200 headphones.

## 2.5 Statistical models

All statistical analyses were done using the R software package (version 3.6.1). We used the `lm` function to fit our standard linear regression models. For our mixed effect linear regression models, we used the “`lme4`” library ([Bates et al., 2014](#)) to fit the models and calculate  $t$ -values, and the “`lmer`” test package ([Kuznetsova et al., 2017](#)) to estimate  $p$ -values and degrees of freedom. The FDR adjusted  $p$ -values were calculated using the “`p.adjust`” function in R. Before model fitting, all continuous values were standardized and all categorical variables were coded as 0 and 1 (i.e., for sound type, music = 0 and voice = 1).

### 3. Results

Recall that our first hypothesis predicted that scream-like music and screams would share a similar roughness level that would be higher than their matched controls. To test our first hypothesis, we used a standard general linear regression analysis. The predicted value was roughness and the predictor value was scream-likeness. We also tested for an interaction effect between sound types and scream-likeness. Finally, another standard general linear regression analysis was used to test for a main effect of scream-likeness on roughness for just the voice (coded as 1) for replicative comparison to the findings of Arnal *et al.* (2015). The results (reported in Table 1) demonstrate a significant main effect between scream-likeness and roughness driven by higher roughness values for scream-like stimuli as compared to non-scream-like stimuli across both sound types ( $p < 0.001$ );  $R^2$  for the model was 0.407, and adjusted  $R^2$  was 0.401. A similar effect was demonstrated in the regression analysis of the vocal stimuli only ( $p < 0.001$ ) replicating the findings of Arnal *et al.* (2015);  $R^2$  for the model was 0.826, and adjusted  $R^2$  was 0.824. Additionally, the results showed a significant interaction effect between sound type and scream-likeness ( $p < 0.001$ ) driven by a more extreme difference in mean roughness values between screaming voices and non-screaming voices compared to scream-like music versus non-scream-like music;  $R^2$  for the model was 0.503, and adjusted  $R^2$  was 0.495. These results are consistent with our hypotheses that roughness levels would be higher in the scream-like category than the non-scream-like category, across both sound types. However, contrary to our hypothesis, screams had a significantly higher mean roughness ( $M = 4.67$ ,  $SD = 0.40$ ) than scream-like music ( $M = 4.10$ ,  $SD = 0.75$ ,  $p < 0.001$ , see Table S3 for all unstandardized means and SDs<sup>1</sup>).

Our second hypothesis was that roughness would correlate negatively with valence ratings and positively with arousal ratings for both music and vocal stimuli, supporting its reputation as an aural cue for danger (Arnal *et al.*, 2015). To test this hypothesis, we used emotion ratings (valence and arousal) as the predicted values for two mixed effects linear regression models. The predictor value was roughness. Once again, participant was included as a random slope. We also tested for a main effect for roughness on emotion ratings for vocal stimuli only for further replicative comparison to the findings of Arnal *et al.* (2015). Additionally, we tested for an interaction effect between sound type (music vs voice) and roughness. The regression analyses results are reported in Table 2. Consistent with our hypothesis, roughness correlated negatively with valence ratings for both musical stimuli [ $\beta = -0.196$ ,  $SE = 0.032$ ,  $t = -6.15$ ,  $p < 10^{-4}$ , BH-adjusted  $p < 10^{-4}$ ] and vocal stimuli [ $\beta = -0.443$ ,  $SE = 0.052$ ,  $t = -8.599$ ,  $p < 10^{-8}$ , BH-adjusted  $p < 10^{-7}$ ]. Also consistent with our third hypothesis, roughness correlated positively with arousal ratings for both musical stimuli [ $\beta = 0.07$ ,  $SE = 0.025$ ,  $t = 2.859$ ,  $p = 0.01$ , BH-adjusted  $p = 0.0134$ ] and vocal stimuli [ $\beta = 0.404$ ,  $SE = 0.057$ ,  $t = 7.144$ ,  $p < 10^{-7}$ , BH-adjusted  $p < 10^{-6}$ ]. Interestingly, there were significant interaction effects between roughness and sound type for both valence [ $\beta = -0.247$ ,  $SE = 0.045$ ,  $t = -5.441$ ,  $p < 10^{-5}$ , BH-adjusted  $p < 10^{-5}$ ] and arousal ratings [ $\beta = 0.334$ ,  $SE = 0.058$ ,  $t = 5.746$ ,  $p < 10^{-5}$ , BH-adjusted  $p < 10^{-5}$ ]. Both regression slopes were steeper for the vocal sound type than for the musical sound type. The interaction effect models can be seen in the scatterplots in Fig. 1(B). In spite of the significant interaction effect, the significant main effects for both valence [ $\beta = -0.330$ ,  $SE = 0.037$ ,  $t = -8.842$ ,  $p < 10^{-8}$ , BH-adjusted  $p < 10^{-7}$ ] and arousal [ $\beta = 0.251$ ,  $SE = 0.034$ ,  $t = 7.342$ ,  $p < 10^{-7}$ , BH-adjusted  $p < 10^{-6}$ ] indicate that the relationship between roughness and emotion ratings extends across sound types, consistent with our hypothesis.

### 4. Discussion and conclusions

This research was inspired by the frequent comparison of scary film music to human screams (Brown, 1982). Our motivating question was whether scary film music mimics an acoustic feature

Table 1. Roughness as predicted by scream-likeness and sound type.

	Roughness (voice only)	Roughness (main effect)	Roughness (interaction)
Intercept	<b>-0.88<sup>a</sup></b> [-1.00, -0.75]	<b>-0.70<sup>a</sup></b> [-0.89, -0.52]	<b>-0.39<sup>a</sup></b> [-0.59, -0.19]
Scream-likeness (screaming = 1)	<b>1.89<sup>a</sup></b> [1.72, 2.06]	<b>1.27<sup>a</sup></b> [1.05, 1.48]	<b>0.65<sup>a</sup></b> [0.37, 0.93]
Sound type (voice = 1)		0.14 [-0.08, 0.35]	<b>-0.48<sup>b</sup></b> [-0.76, -0.21]
Scream-likeness $\times$ sound type			<b>1.24<sup>a</sup></b> [0.85, 1.64]
<i>N</i>	100	200	200
$R^2/R^2_{adjusted}$	0.826/0.824	0.407/0.401	0.503/0.495

<sup>a</sup> $p < 0.001$ .

<sup>b</sup> $p < 0.01$ .

Table 2. Valence and arousal ratings as predicted by roughness and sound type (df = 19, bold results are statistically significant at adjusted  $p < 0.05$ ).

	Est.	SE	<i>t</i>	Unadj. <i>p</i>	BH-adj. <i>p</i>
Valence (voice only)					
Intercept	<b>−0.351</b>	<b>0.086</b>	<b>−4.085</b>	<b>6.31E−04</b>	<b>6.31E−04</b>
Roughness	<b>−0.443</b>	<b>0.052</b>	<b>−8.599</b>	<b>5.65E−08</b>	<b>1.13E−07</b>
Valence (main effect)					
Intercept	<b>0.358</b>	<b>0.079</b>	<b>4.511</b>	<b>2.39E−04</b>	<b>2.39E−04</b>
Roughness	<b>−0.330</b>	<b>0.037</b>	<b>−8.842</b>	<b>3.68E−08</b>	<b>1.10E−07</b>
Sound type (voice = 1)	<b>−0.717</b>	<b>0.122</b>	<b>−5.884</b>	<b>1.15E−05</b>	<b>1.73E−05</b>
Valence (interaction)					
Intercept	<b>0.368</b>	<b>0.079</b>	<b>4.642</b>	<b>1.78E−04</b>	<b>1.78E−04</b>
Roughness	<b>−0.196</b>	<b>0.032</b>	<b>−6.150</b>	<b>6.54E−06</b>	<b>2.22E−05</b>
Sound type (voice = 1)	<b>−0.718</b>	<b>0.122</b>	<b>−5.900</b>	<b>1.11E−05</b>	<b>2.22E−05</b>
Roughness × sound type	<b>−0.247</b>	<b>0.045</b>	<b>−5.441</b>	<b>3.00E−05</b>	<b>4.00E−05</b>
Arousal (voice only)					
Intercept	0.209	0.107	1.956	0.0654	0.0654
Roughness	<b>0.404</b>	<b>0.057</b>	<b>7.144</b>	<b>8.63E−07</b>	<b>1.73E−06</b>
Arousal (main effect)					
Intercept	−0.220	0.114	−1.931	0.0685	0.0685
Roughness	<b>0.251</b>	<b>0.034</b>	<b>7.342</b>	<b>5.86E−07</b>	<b>1.76E−06</b>
Sound type (voice = 1)	<b>0.439</b>	<b>0.108</b>	<b>4.079</b>	<b>6.39E−04</b>	<b>9.59E−04</b>
Arousal (interaction)					
Intercept	−0.232	0.115	−2.018	0.0579	0.0579
Roughness	<b>0.070</b>	<b>0.025</b>	<b>2.859</b>	<b>0.0100</b>	<b>0.0134</b>
Sound type (voice = 1)	<b>0.441</b>	<b>0.108</b>	<b>4.092</b>	<b>0.0001</b>	<b>0.0012</b>
Roughness × sound type	<b>0.334</b>	<b>0.058</b>	<b>5.746</b>	<b>1.55E−05</b>	<b>6.20E−05</b>

unique to human screams (roughness) (Arnal *et al.*, 2015) in order to scare the viewers. In order to address this question, we calculated roughness levels (specifically, the average MPS power in the 30–150 Hz region) (Arnal *et al.*, 2015) for four groups of audio recordings: screams, non-screaming vocalizations, scream-like music, and non-scream-like music. We also ran a behavioral study where 20 participants rated the arousal and valence of each audio file.

Consistent with our hypotheses, we found that both screams and scream-like music exhibited a higher level of roughness and were rated as having a more negative valence and a higher arousal level than their non-screaming counterparts. However, contrary to our hypotheses, screams had a higher roughness level than scream-like music. Overall, the results demonstrated a greater difference in roughness levels and emotion ratings between the vocal stimuli than between the musical stimuli. These results suggest that while scream-like music does seem to sound like and be perceived similarly to human screams, the musical rendition is still a muted version of the real thing and therefore may not provoke as potent of a reaction. This finding is notably in opposition to the super-expressive voice theory (Juslin and Västfjäll, 2008) that music is capable of amplifying vocal affective behaviors beyond the capability of the vocal system. Perhaps screams are an exception to this theory. Overall, the results suggest that roughness can effectively translate from a vocal cue for danger into a musical cue for danger. It is therefore reasonable to suggest that scream-like music might scare viewers in part because it is evocative of a human scream, a naturally alarming sound.

It is important to note that our results may have been biased by the non-blind selection of the film music excerpts. This work corroborates the findings of several recent papers investigating roughness in music (Arnal *et al.*, 2015; Belin and Zatorre, 2015; Blumstein *et al.*, 2010; Blumstein *et al.*, 2012; Liuni *et al.*, 2020; Ollivier *et al.*, 2019) while being the first to use the MPS to investigate the presence of roughness in scary film music. Future work might test similar hypotheses by directly manipulating roughness cues in music [i.e., Anikin (2020)], by investigating other features associated with vocal roughness such as jitter, shimmer, and HNR (Liuni *et al.*, 2020), or by investigating whether our findings could be extended to other music found to induce negative emotions, such as metal music for non-fans.

## Acknowledgments

C.T. received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement (No. 835682). S.F. received funding from

Swiss National Science Foundation (Grants Nos. SNSF PP00P1\_157409/1 and PP00P1\_183711/1). The authors thank Lawrence Feth for guidance regarding the acoustic analyses and to David Huron for valuable feedback on the project. Finally, the authors thank Arkady Konovalov for helpful input regarding the statistical analyses.

## References and links

<sup>1</sup>See supplementary material at <https://doi.org/10.1121/10.0001459> for Tables S1–S3.

- Anikin, A. (2020). “The perceptual effects of manipulating nonlinear phenomena in synthetic nonverbal vocalizations,” *Bioacoustics* **29**(2), 226–247.
- Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A.-L., and Poeppel, D. (2015). “Human screams occupy a privileged niche in the communication soundscape,” *Curr. Biol.* **25**(15), 2051–2056.
- Arnal, L. H., Kleinschmidt, A., Spinelli, L., Giraud, A.-L., and Mégevand, P. (2019). “The rough sound of salience enhances aversion through neural synchronisation,” *Nat. Commun.* **10**(1), 1–12.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). “Fitting linear mixed-effects models using lme4,” *arXiv:1406.5823*.
- Belin, P., and Zatorre, R. J. (2015). “Neurobiology: Sounding the alarm,” *Curr. Biol.* **25**(18), R805–R806.
- Blumstein, D. T., Bryant, G. A., and Kaye, P. (2012). “The sound of arousal in music is context-dependent,” *Biol. Lett.* **8**(5), 744–747.
- Blumstein, D. T., Davitian, R., and Kaye, P. D. (2010). “Do film soundtracks contain nonlinear analogues to influence emotion?,” *Biol. Lett.* **6**(6), 751–754.
- Brown, R. S. (1982). “Herrmann, Hitchcock, and the music of the irrational,” *Cinema J.* **21**(2), 14–49.
- Bryant, G. A. (2013). “Animal signals and emotion in music: Coordinating affect across groups,” *Front. Psychol.* **4**, 990.
- Ehret, G. (2006). “Common rules of communication sound perception,” in *Behavior and Neurodynamics for Auditory Communication*, edited by J. S. Kanwal and G. Ehret (Cambridge University Press, Cambridge), pp. 85–114.
- Elliott, T. M., and Theunissen, F. E. (2009). “The modulation transfer function for speech intelligibility,” *PLoS Comput. Biol.* **5**(3), 1–14.
- Engelberg, J. W., and Gouzoules, H. (2018). “The credibility of acted screams: Implications for emotional communication research,” *Quart. J. Exp. Psychol.* **72**, 1889–1902.
- Huron, D. (2015). “Cues and signals: An ethological approach to music-related emotion,” *Signata* **6**, 331–351.
- Huron, D., and Trevor, C. (2017). “Are stopped strings preferred in sad music?,” *Empirical Musicol. Rev.* **11**(2), 261–269.
- Juslin, P. N., and Laukka, P. (2003). “Communication of emotions in vocal expression and music performance: Different channels, same code?,” *Psychol. Bull.* **129**(5), 770–814.
- Juslin, P. N., and Västfjäll, D. (2008). “Emotional responses to music: The need to consider underlying mechanisms,” *Behav. Brain Sci.* **31**(05), 559–575.
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). “lmerTest package: Tests in linear mixed effects models,” *J. Stat. Softw.* **82**(13), 1–26.
- Li, T., Horta, M., Mascaro, J. S., Bijanki, K., Arnal, L. H., Adams, M., Barr, R. G., and Rilling, J. K. (2018). “Explaining individual variation in paternal brain responses to infant cries,” *Physiol. Behav.* **193**, 43–54.
- Liuni, M., Ponsot, E., Bryant, G. A., and Aucouturier, J. J. (2020). “Sound context modulates perceived vocal emotion,” *Behav. Proc.* **172**, 104042.
- Lorenz, K. (1939). “Vergleichende verhaltensforschung,” (“Comparative behavioral research”), *Zool. Anz.* **12**, 69–102.
- Mirka, D. (2014). *The Oxford Handbook of Topic Theory* (Oxford University Press, Oxford).
- Ohala, J. J. (1996). “Ethological theory and the expression of emotion in the voice,” in *Proceeding of Fourth International Conference on Spoken Language Processing, ICSLP '96*, Vol. 3, pp. 1812–1815.
- Ollivier, R., Goupil, L., Liuni, M., and Aucouturier, J.-J. (2019). “Enjoy the violence: Is appreciation for extreme music the result of cognitive control over the threat response system?,” *Music Percept.: Interdisc. J.* **37**(2), 95–110.
- Schwartz, J. W., Engelberg, J. W., and Gouzoules, H. (2019). “What is a scream? Acoustic characteristics of a human call type,” *J. Acoust. Soc. Am.* **145**(3), 1776–1776.
- Terhardt, E. (1974). “On the perception of periodic sound fluctuations (roughness),” *Acta Acust. Acust.* **30**(4), 201–213.
- Thompson, W. F., Geeves, A. M., and Olsen, K. N. (2019). “Who enjoys listening to violent music and why?,” *Psychol. Pop. Media Culture* **8**(3), 218–232.
- Trevor, C., Arnal, L., and Frühholz, S. (2020). <https://osf.io/7d2cy/> (Last viewed 06/17/2020.)
- Trevor, C., and Huron, D. (2019). “Are humoresques humorous? On the similarity between laughter and staccato,” *Emp. Musicol. Rev.* **13**(1-2), 66–77.
- Vassilakis, P. N., and Kendall, R. A. (2010). “Psychoacoustic and cognitive aspects of auditory roughness: Definitions, models, and applications,” in *Human Vision and Electronic Imaging XV* (International Society for Optics and Photonics, Bellingham, WA), Vol. 7527, p. 75270O.
- Warrenburg, L. A. (2019). “Comparing musical and psychological emotion theories,” *Psychomusic.: Music Mind Brain* **30**(1), 1–19.